ISSN: 1934--9955 www.ijise.net
Vol-20 Issue-01 Mar 2025

SENTIMENT BASED DRUG RECOMMENDATION USING MACHINE LEARNING

[1]N.Lavanya, [2]Rosline, [3]M.Rohith Reddy, [4]M. Santhosh naik,[5] N. Srikanth,
[1,2,3,4] U.G.Scholor, Department of ECE, Sri Indu College Of Engineering & Technology, Ibrahimpatnam, Hyderabad.
[5]Guide Research, Department of ECE, Sri Indu College Of Engineering & Technology, Ibrahimpatnam, Hyderabad.

ABSTRACT:

Early detection of Alzheimer's disease (AD) is based on the categorization of characteristics retrieved from brain scans, which plays a significant role in preventing and treating the illness. The characteristics must properly reflect key AD-related variations in physical brain structures such ventricles size, hippocampus shape, cortical thickness, and brain volume. This research suggests using a deep 3D convolutional neural network (3D-CNN) to predict Alzheimer's disease using generic features that capture AD biomarkers and can adapt to different domain datasets. The 3D-CNN is based on a pre-trained 3D convolutional autoencoder for capturing anatomical shape changes in structural brain MRI data. For each task-specific AD classification, the fully linked higher layers of the 3D-CNN are fine-tuned. ADNI MRI experiments

INTRODUCTION:

Alzheimer's disease (AD), a degenerative brain ailment and the most prevalent form of dementia in older people, involves nerve cell death and tissue loss throughout the brain, resulting in a substantial reduction in brain volume over time and impacting most of its functions [1]. By 2050, one out of every 85 individuals will have Alzheimer's disease [2]. Because the expense of caring for Alzheimer's patients is predicted to skyrocket, having a computer-aided approach for early and accurate AD detection becomes increasingly important [3]. The goal of this study is to create an extensible deep learning-based system for early Alzheimer's disease diagnosis. Deep learning assists in problem solving.

PRIOR WORK:

The sMRI parameters of voxel-wise, cortical thickness, and hippocampus shape volume are employed to diagnose AD [3]. After co-aligning (registering) all the brain imaging data, voxel-wise features are retrieved to connect each brain voxel with a vector (signature) of several scalar measures. Kloppel et al. [15] employed grey matter (GM) voxels as characteristics and trained an SVM to distinguish between AD and NC patients. In [16], the brain volume is divided into GM, white matter (WM), and CSF sections, with voxel-wise densities calculated and each voxel associated with a vector of GM, WM, and CSF densities for classification. Lerch et al. [17] segmented the recorded brain MRI into the GM, WM, and M.

MODEL:

A source-domain-trained 3D-CAE extracts features from a brain MRI, and a highly supervised target-domain-adaptable 3D-CNN conducts task-specific classification. Sections 3.1 and 3.3 describe the 3D-CAE architecture and the AD diagnostic framework employing the DSA-3D-CNN, respectively.

Convolutional Autoencoder (3D-CAE):

The traditional unsupervised autoencoder combines data encoding and decoding to extract a few co aligned scalar feature maps from a group of input 3D pictures containing scalar or vectorial voxel-wise signals. The input picture is encoded in the hidden layer by mapping each fixed voxel neighbourhood to a vectorial feature space and then reconstructed to the original image space in the output layer. To decrease the reconstruction error, back-propagation and limitations on feature space attributes are used in autoencoder training to extract features that represent typical patterns of input data fluctuations. Using vectorial voxel-wise signals to extract global features from 3D pictures is computationally intensive and necessitates a large training data set. This is attributed to the rapidly increasing number of parameters to be considered.

$$h_{i:j:k} = f\left(\mathbf{W}_k * \mathbf{x}_{i:\text{neib}} + b_{j:k}\right) \qquad (1)$$

The latter function is chosen from a wide range of restricting differentiable functions, including the sigmoid and others.

$f(u) = (1 + \exp(-u))^{-1}$ and a linear rectified unit (ReLU),

$f(u) = \max(0, u)$ [35]. Because the J-vectorial is present in the 3D picture x in Eq. (1)

Weights Wk form a 3D moving-window filter for each voxel inside voxel-wise signals, which convolve the union of J dimensional signal spaces. To make things easier, we'll use

$$\mathbf{h}_k = \mathbb{T}\left(\mathbf{x} : \mathbf{W}_k, \mathbf{b}_k, f(\cdot)\right)$$

indicate the whole encoding of the input 3D picture with J vectorial voxel-wise signals with the k-th 3D feature map, hk, so that its scalar components are produced with Eq. (1) for a given voxel neighbourhood using the weights Wk and bias vectors bk. The equivalent inverse transformation, Tinv(...), decodes or reconstructs the initial 3D image:

$$\widehat{\mathbf{x}} = \sum_{k=1}^{K} \underbrace{\mathbf{T}_{\text{inv}}\left(\mathbf{h}_k : \mathbf{P}_k, \mathbf{b}_{\text{inv}:k}, g(\cdot)\right)}_{\mathbf{a}_k} \qquad (2)$$

Given L encoding layers, each layer l generates an output feature image, $\mathbf{h}_{(l)} = [\mathbf{h}_{(l):k} : k = 1, \ldots, K_l]$, with K_l-vectorial voxel-wise features and receives the preceding output, $\mathbf{h}_{(l-1)} = [\mathbf{h}_{(l-1):k} : k = 1, \ldots, K_{l-1}]$ as the input image (i.e., $\mathbf{h}_{(0)} = \mathbf{x}$).

The 3D-CAE of Eqs. (1) and (2) is trained by minimizing the mean squared reconstruction error for T; $T \geq 1$, given training input images, $\mathbf{x}^{[t]}$; $t = 1, \ldots, T$,

$$E(\boldsymbol{\theta}) = \frac{1}{T} \sum_{t=1}^{T} \| \widehat{\mathbf{x}}^{[t]} - \mathbf{x}^{[t]} \|_2^2 \qquad (3)$$

where $\boldsymbol{\theta} = [\mathbf{W}_k; \mathbf{P}_k; \mathbf{b}_k; \mathbf{b}_{\text{inv}:k} : k = 1, \ldots, K]$, and $\| \ldots \|_2^2$ denote all free parameters and the average vectorial ℓ_2-norm over the T training images, respectively. To reduce the

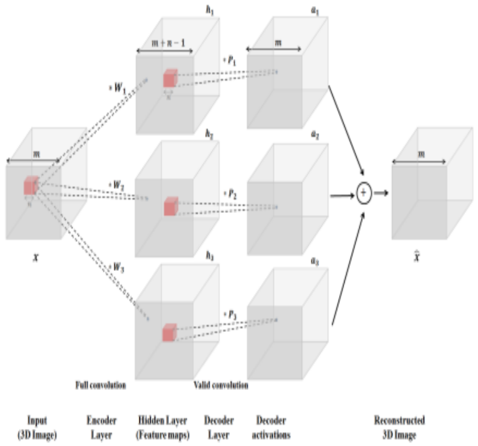
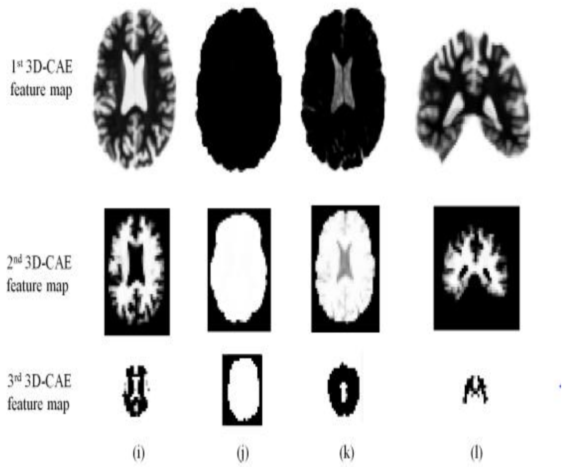


Fig. 1: CAD Dementia brains MRI at three layers of the stacked 3D-CAE: cortex thickness and volume I brain size (j), ventricle size (k), and hippocampal size (l) (l). To minimise their size and identify higher-level features, the feature maps are down-sampled at each layer using max-pooling.



The decoding and encoding weights Pk and Wk were connected by flipping over all their dimensions, as proposed in [32]. Using the stochastic gradient descent search and error back-propagation, the cost function Eq. (3) was minimised in the parameter space. The feature maps, h(i), are down-sampled using max-pooling, i.e. extracting the maximum value of non-overlapping sub-regions, to provide translational invariance. The max-pooling result is utilised to train the upper layer CAE for entangling form changes in lower-level feature maps of decreasing size, as seen in Fig. 1. (a). At each level of its hierarchy, stacking the encoding 3D-CAE layers (abbreviated 3D-CAES below) reduces the size of the feature map by half [32].

Transfer Learning and Domain Adaptation:

A large training set of labelled data is generally required for supervised learning of a classifier to get decent results. If this collection is in theory too small, transfer learning can be used to include extra information gained during the construction of a comparable classifier. Initial weights learnt for performing comparable tasks [36–39] might be used in the goal classifier based on a deep CNN. We concentrate on domain adaptation [40–42], also known as source-to-target adaptation, which occurs when a classifier is adapted to target data after being trained on source data. Unlike traditional guided learning, which involves training a classifier from start and minimising a total quantitative loss from mistakes on the training data, domain adaptation involves minimising the same loss throughout the target domain.

Deeply Supervised Adaptive 3D-CNN (DSA-3DCNN:

While the bottom layers of a goal predictive 3D-CNN extract generic characteristics, the top levels must make task specific categorization easier utilising those features [6]. The proposed classifier collects generic characteristics using a stack of locally connected lower convolutional layers while fine-tuning parameters of the fully connected top layers for task-specific purposes. Pre-training, initial training of the lower convolutional layers, and task-specific fine-tuning comprise the proposed hierarchical 3D CNN's training. The convolutional layers for generic feature extraction are created as a stack of 3D-CAEs that have been pre-trained in the source domain during the pre-training step. Then these layers are initialised by encoding the 3D-CAE weights [5], followed by fine-tuning of the higher fully loaded layers using deep supervision [14].

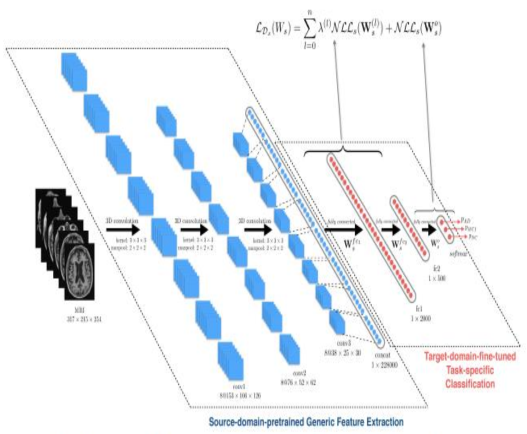


Fig. 2: Architecture of the deeply supervised and adaptable 3D CNN (DSA-3D-CNN) for AD diagnosis.

pictures from the network's bottom pre-trained layer The ReLU activation functions at each inner layer and the fully linked higher layers with a softmax top-most output layer (Fig. 2) are used in our implementation of the 3D-CNN to predict whether an input brain sMRI belongs to the AD, MCI, or NC groups. Deep supervision of this 3D-higher CNN's layers increased its task-specific performance [14,45]. It optimises a weighted sum of the like log-likelihood-based sepa rate losses, dependent on the weights for each individual fully connected upper layer, plus the like loss of the top-most layer conducting the softmax transformation of its convolved and biassed inputs, enabling task-specific fine-tuning. A classifier's discriminative skills are

EXPERIMENTS:

The proposed DSA-3D-CNN for AD diagnosis was validated on 30 CADDementia subjects as the source domain and 210 ADNI subjects as the target domain (demographic information in Table 1) for five classification tasks: four binary ones (AD vs. NC, AD+MCI vs. NC, AD vs. MCI, MCI vs NC), and the ternary classification (AD vs. MCI vs. NC). Each test's classification accuracy was assessed using ten-fold cross-validation. The deep CNN deployed for our research on Amazon EC2 g2.8xlarge instances with GPU GRID K520 was built using the Theano library [47].

Table 1: ADNI database (STD = standard deviation), demographic data for 210 participants from the target domain.

Diagnosis	AD	MCI	NC
Number of subjects	70	70	70
Male / Female	36 / 34	50 / 20	37 / 33
Age (mean$_{\pm STD}$)	$75.0_{\pm 7.9}$	$75.9_{\pm 7.7}$	$74.6_{\pm 6.1}$

Generic and task-specific feature evaluation:

Special 2D projections of the retrieved features in Fig. 1(b) demonstrate the proposed DSA-3D-generalization CNN's and adaption capability (Fig. 2). In Fig. 1(b), selected slices of the three feature maps from each layer of

International journal of imaging science and engineering
ISSN: 1934--9955 www.ijise.net
Vol-20 Issue-01 Mar 2025

our stacked 3D-CAE (abbreviated 3D-CAES below) reveal that the learned generic convolutional filters can capture variables related to AD biomarkers, such as ventricle size, cortical thickness, and hippocampus model. The pre-trained 3D-CAES created these feature maps for the CADementia database. The first layer of the 3D-CAES extracts the cortical thickness as a discriminative AD characteristic based on these projections.

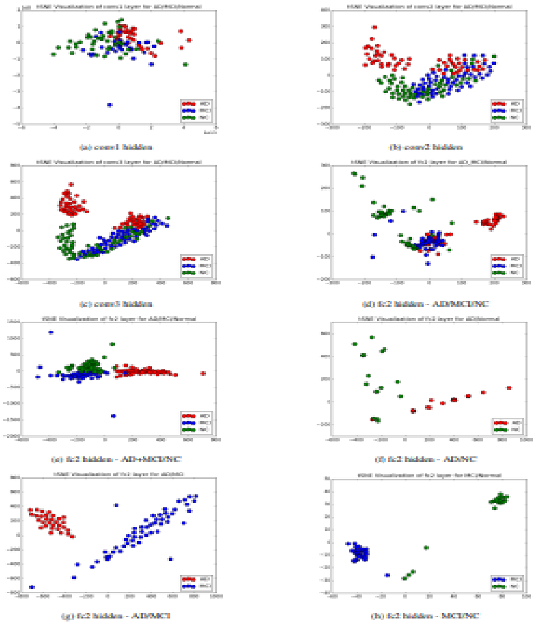


Table 2: Performance of the proposed classifier on the target domain (ADNI) for a certain cross-validation fold (see Section 4.2 and Eq (4).

	AD / MCI / NC			AD+MCI / NC			AD / NC			AD / MCI			MCI / NC		
Class	PPVr	SEN	F1	PPV	SEN	F1	PPV	SEN	F1	PPV	SEN	F1	PPV	SEN	F1
AD	1.00	1.00	1.00	-	-	-	0.88	1.00	0.94	1.00	1.00	1.00	-	-	-
MCI	0.60	0.80	0.69	-	-	-	-	-	-	1.00	1.00	1.00	0.92	0.97	0.94
AD+MCI	-	-	-	0.94	0.97	0.95	-	-	-	-	-	-	-	-	-
NC	0.70	0.47	0.56	0.93	0.87	0.90	1.00	0.87	0.93	-	-	-	0.97	0.91	0.94
Mean	0.77	0.76	0.75	0.93	0.93	0.93	0.94	0.93	0.93	1.00	1.00	1.00	0.95	0.94	0.94

The ensuing layers reflect the brain size (related to the patient gender), ventricle size, and hippocampus model of AD. Each 3D-CAES layer combines the retrieved lower-layer feature maps to train the upper level to describe the anatomical changes in the brain sMRI in more detail. At the next levels, both the ventricle size and the cortical thickness characteristics are merged to get conceptually higher level features. Projection capabilities of the recovered higher-layer features to differentiate the AD, MCI, and NC brain sMRIs in the low-dimensional feature space are shown in Fig. 3. Projected manifold distributions of the training ADNI sMRI over the hidden layers of our DSA 3D-CNN are shown in Fig. 3.

Classification performance evaluation:

Eight assessment criteria were used to evaluate and compare the proposed DSA-3D-CNN classifier's performance for each of the tasks described in Section 4.1. For a given set of data items, let TP, TN, FP, and FN indicate the number of true positive, true negative, false positive, and false negative classification findings. The following measures [49] are used to evaluate performance: accuracy (ACC); sensitivity (SEN), or recall; specificity (SPE); balanced accuracy (BAC); positive predictive value (PPV), or precision; negative predictive value (NPV), and F1-score.

International journal of imaging science and engineering

ISSN: 1934--9955 www.ijise.net
Vol-20 Issue-01 Mar 2025

$$\begin{aligned} &\mathrm{ACC}=\frac{\mathrm{TP}+\mathrm{TN}}{\mathrm{TP}+\mathrm{TN}+\mathrm{FP}+\mathrm{FN}}; && \mathrm{F1}=\frac{2\cdot\mathrm{TP}}{2\cdot\mathrm{TP}+\mathrm{FP}+\mathrm{FN}}; \\ &\mathrm{SEN}=\frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FN}}; && \mathrm{SPE}=\frac{\mathrm{TN}}{\mathrm{TN}+\mathrm{FP}}; \\ &\mathrm{PPV}=\frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FP}}; && \mathrm{NPV}=\frac{\mathrm{TN}}{\mathrm{TN}+\mathrm{FN}}; \\ &\mathrm{BAC}=\tfrac{1}{2}(\mathrm{SEN}+\mathrm{SPE}) \end{aligned} \quad (4)$$

Furthermore, the area under the ROC curve is used to evaluate the classifier's performance after constructing a receiver operating characteristic (ROC) (AUC). Table 2 shows the performance of our DSA 3D-CNN classifier for five different classification tasks and a specific cross-validation fold. The ROCs / AUCs of these tests in Fig. 5 and the means and standard deviations of all the met rics of Eq. (4) in Table 3 suggest that the proposed task-specific DSA-3D-NCC classifier's AD predictions are robust and consistent. Table 4 compares its accuracy (ACC) to that of seven other well-known techniques that employ the same or even more inputs (imaging modalities).

Our classifier's ten-fold cross-validation average results are shown in Table 4. When compared to existing techniques in all five task-specific scenarios, the suggested DSA-3D-CNN beats them. Despite the fact that just one imaging modality (sMRI) was used and no previous skull-stripping was performed, this result was obtained.

CONCLUSION:

This study presented a DSA-3D-CNN classifier that outperforms multiple existing state-of-the-art predictors in predicting AD on structural brain MRI data. With three layered 3D CAE networks pre-trained on the CADDementia dataset, the transfer learning approach is employed to improve generality of the features collecting AD biomarkers. The characteristics are then retrieved and employed in the bottom layers of a 3D CNN network to identify AD biomarkers. The lowest layer is then piled with three completely linkelayers on top of it.

Table 3: Performance of the proposed DSA-3D-NCC classifier on target domain (*ADNI*) [$\mathrm{mean}_{\mathrm{STD}}$,%].

Task	Performance metrics (Section 4.2:							
	ACC	SEN	SPE	BAC	PPV	NPV	AUC	F1-score
AD / MCI / NC	$\mathbf{94.8_{2.6}}$	–	–	–	–	–	–	–
AD+MCI / NC	$\mathbf{95.7_{3.1}}$	$94.8_{4.1}$	$97.2_{3.8}$	$96.0_{2.9}$	$98.4_{2.2}$	$91.0_{6.8}$	$96.1_{2.9}$	$93.9_{4.4}$
AD / NC	$\mathbf{99.3_{1.6}}$	100_0	$98.6_{3.1}$	$99.3_{1.6}$	$98.6_{3.1}$	100_0	$99.3_{2.0}$	$99.4_{1.3}$
AD / MCI	$\mathbf{100_0}$	100_0	100_0	100_0	100_0	100_0	100_0	100_0
MCI / NC	$\mathbf{94.2_{2.0}}$	$97.1_{5.7}$	$91.4_{4.0}$	$94.3_{2.0}$	$91.9_{4.3}$	$97.1_{4.5}$	$97.1_{2.0}$	$94.4_{1.7}$

Table 4: Comparative performance (ACC,%) of the classifier vs. seven competitors on *ADNI* dataset (n/a – non-available).

Approach	Modalities	Task-specific classification [$\mathrm{mean}_{\mathrm{STD}}$,%].				
		AD/MCI/NC	AD+MCI/NC	AD/NC	AD/MCI	MCI/NC
Gupta et al. [20]	MRI	$85.0_{n/a}$	n/a	$94.7_{n/a}$	$88.1_{n/a}$	86.3_{rm}
Suk et al. [22]	PET+MRI+CSF	n/a	n/a	$95.9_{1.1}$	n/a	$85.0_{1.2}$
Suk et al. [23]	PET+MRI	n/a	n/a	$95.4_{5.2}$	n/a	$85.7_{5.2}$
Zhu et al. [25]	PET+MRI+CSF	n/a	n/a	$95.9_{n/a}$	n/a	$82.0_{n/a}$
Zu et al. [27]	PET+MRI	n/a	n/a	$96.0_{n/a}$	n/a	$80.3_{n/a}$
Liu et al. [24]	PET+MRI	$53.8_{4.8}$	n/a	$91.4_{5.6}$	n/a	$82.1_{4.9}$
Payan et al. [28]	MRI	$89.4_{n/a}$	n/a	$95.39_{n/a}$	$86.8_{n/a}$	$92.1_{n/a}$
Liu et al. [19]	MRI	n/a	n/a	$93.8_{n/a}$	n/a	$89.1_{n/a}$
Li et al. [26]	PET+MRI+CSF	n/a	n/a	$91.4_{1.8}$	$70.1_{2.3}$	$77.4_{1.7}$
Our DSA-3D-CNN	MRI	$\mathbf{94.8_{2.6}}$	$\mathbf{95.7_{3.1}}$	$\mathbf{99.3_{1.4}}$	$\mathbf{100_0}$	$\mathbf{94.2_{2.0}}$

ers to do AD classification on 210 ADNI dataset participants. To increase topic discrimination, a discriminative loss function was used to each fully connected layer in addition to the output classification layers to improve classification performance. The findings show that hierarchical feature extraction improved in 3D-CNN hidden layers, distinguishing between AD, MCI, and NC individuals. We looked at seven different categorizationmeasures.

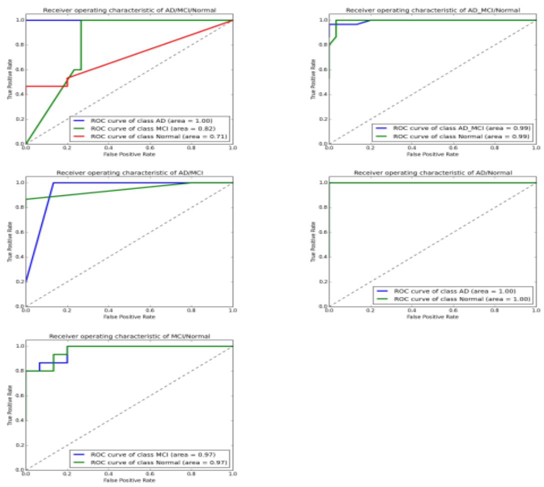


Fig. 5: ROCs and AUC performance scores for the DSA-3D-CNN classifier were compared to state-of-the-art models after fine-tuning to the specific task of distinguishing between (left-to-right) AD / MCI / NC; AD+MCI / NC; AD / NC; AD / MCI, and MCI / NC subjects on target domain (ADNI) using ten-fold crossvalidation. The findings show that the proposed DSA-3D-CNN outperforms others.

ACKNOWLEDGMENT:The Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904) and DOD ADNI financed the data gathering and sharing for this investigation (Department of Defense award number W81XWH-12- 2-0012). The National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and the following generous donors support ADNI: AbbVie, Alzheimer's Association, Alzheimer's Drug Discovery Foundation; Araclon Biotech; BioClinica, Inc.; Biogen; Bristol-Myers Squibb Company; CereSpir, Inc.; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; Fujirebio; GE Healthcare; IX ICO Ltd

REFERENCES:

[1] G. McKhann, D. Drachman, M. Folstein, R. Katzman, D. Price, and E. Stadlan, “Clinical diagnosis of Alzheimer’s disease: Report of the NINCDSADRDA Work Group* under the auspices of Department of Health and Human Services Task Force on Alzheimer’s Disease,” Neurology, vol. 34, no. 7, p. 939, 1984.

[2] A. Association et al., “2014 Alzheimer’s disease facts and figures,” Alzheimer’s & Dementia, vol. 10, no. 2, pp. e47–e92, 2014.

[3] E. Bron, M. Smits, W. van der Flier et al., “Standardized evaluation of algorithms for computer-aided diagnosis of dementia based on structural MRI: The CADDementia challenge,” NeuroImage, 2015.

[4] S. Plis, D. Hjelm, R. Salakhutdinov, E. Allen, H. Bockholt, J. Long, H. Johnson, J. Paulsen, J. Turner, and V. Calhoun, “Deep learning for neuroimaging: a validation study,” Frontiers in Neuroscience, vol. 8, 2014.

[5] T. Chen, I. Goodfellow, and J. Shlens, “Net2Net: Accelerating learning via knowledge transfer,” arXiv:1511.05641 [cs.LG], 2015.

[6] M. Long and J. Wang, “Learning transferable features with deep adaptation networks,” arXiv:1502.02791 [cs.LG], 2015.

[7] J. Yosinski, J. Clune, Y. Bengio, and H. Lipson, "How transferable are features in deep neural networks?" in Advances in Neural Information Processing Systems, 2014, pp. 3320–3328.

[8] C. Jack, M. Albert, D. Knopman, G. McKhann, R. Sperling, M. Carrillo, B. Thies, and C. Phelps, "Introduction to the recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease," Alzheimer's & Dementia, vol. 7, no. 3, pp. 257–262, 2011.

[9] G. McKhann, D. Knopman, H. Chertkow, B. Hyman, C. Jack, C. Kawas, W. Klunk, W. Koroshetz, J. Manly, R. Mayeux et al., "The diagnosis of dementia due to Alzheimers disease: Recommendations from the National Institute on Aging-Alzheimers Association workgroups on diagnostic guidelines for Alzheimer's disease," Alzheimer's & Dementia, vol. 7, no. 3, pp. 263– 269, 2011.

[10] R. Cuingnet, E. Gerardin, J. Tessieras, G. Auzias, S. Lehericy, M. Habert, M. Chupin, H. Benali, O. Col- ´ liot, A. D. N. Initiative et al., "Automatic classification of patients with Alzheimer's disease from structural MRI: a comparison of ten methods using the ADNI database," NeuroImage, vol. 56, no. 2, pp. 766–781, 2011.

[11] F. Falahati, E. Westman, and A. Simmons, "Multivariate Data Analysis and Machine Learning in Alzheimer's Disease with a Focus on Structural Magnetic Resonance Imaging," Journal of Alzheimer's Disease, vol. 41, no. 3, pp. 685–708, 2014.

[12] M. Sabuncu and E. Konukoglu, "Clinical Prediction from Structural Brain MRI Scans: A Large-Scale Empirical Study," Neuroinformatics, vol.

13, no. 1, pp. 31– 46, 2015. [13] C. Jack, D. Knopman, W. Jagust, R. Petersen, M. Weiner, P. Aisen, L. Shaw, P. Vemuri, H. Wiste, S. Weigand et al., "Tracking pathophysiological processes in Alzheimer's disease: an updated hypothetical model of dynamic biomarkers," The Lancet Neurology, vol. 12, no. 2, pp. 207–216, 2013. [14] C.-Y. Lee, S. Xie, P. Gallagher, Z. Zhang, and Z. Tu, "Deeply-supervised nets," arXiv:1409.5185 [stat.ML, cs.LG, cs.NE], 2014. [15] S. Kloppel, C. Stonnington, C. Chu, B. Draganski, ¨ R. Scahill, J. Rohrer, N. Fox, C. Jack, J. Ashburner, and R. Frackowiak, "Automatic classification of MR scans in Alzheimer's disease," Brain, vol. 131, no. 3, pp. 681– 689, 2008. [16] Y. Fan, D. Shen, R. Gur, R. Gur, and C. Davatzikos, "COMPARE: classification of morphological patterns using adaptive regional elements," IEEE Trans. Med. Imag., vol. 26, no. 1, pp. 93–105, 2007. [17] J. Lerch, J. Pruessner, A. Zijdenbos, D. Collins, S. Teipel, H. Hampel, and A. Evans, "Automated cortical thickness measurements from MRI can accurately separate Alzheimer's patients from normal elderly controls," Neurobiology of Aging, vol. 29, no. 1, pp. 23–30, 2008. [18] E. Gerardin, G. Chetelat, M. Chupin, R. Cuingnet, ´ B. Desgranges, H. Kim, M. Niethammer, B. Dubois, S. Lehericy, L. Garnero ´ et al., "Multidimensional classification of hippocampal shape features discriminates alzheimer's disease and mild cognitive impairment from normal aging," NeuroImage, vol. 47, no. 4, pp. 1476– 1486, 2009. [19] M. Liu, D. Zhang, E. Adeli-Mosabbeb, and D. Shen, "Inherent structure based multi-view learning with multi-template feature representation for Alzheimer's disease diagnosis," IEEE Trans. Biomed. Eng., vol. PP, no. 99, pp. 1–1, 2015. [20] A. Gupta, M. Ayhan, and A. Maida, "Natural image bases to represent neuroimaging data," in Proceedings of the 30th International Conference on Machine Learning (ICML-13), vol. 28, no. 3. JMLR Workshop and Conference Proceedings, May 2013, pp. 987–994. [21] A. Ng, "Sparse autoencoder," in CS294A Lecture notes. URL https://web.stanford.edu/class/cs294a/ sparseAutoencoder 2011new.pdf: Stanford University, 2011. [22] H.-I. Suk and D. Shen, "Deep learning-based feature representation for ad/mci classification," in Proceedings of the Medical Image Computing and ComputerAssisted Intervention–MICCAI 2013. Springer, 2013, pp. 583–590. [23] H.-I. Suk, S.-W. Lee, D. Shen, A. D. N. Initiative et al., "Hierarchical feature representation and multimodal fusion with deep learning for AD/MCI diagnosis," NeuroImage, vol. 101, pp. 569–582, 2014. [24] S. Liu, S. Liu, W. Cai, H. Che, S. Pujol, R. Kikinis, D. Feng, and M. Fulham, "Multimodal neuroimaging feature learning for multiclass diagnosis of Alzheimer's disease," IEEE Trans. Biomed. Eng., vol. 62, no. 4, pp. 1132–1140, 2015. [25] X. Zhu, H.-I. Suk, and D. Shen, "A novel matrixsimilarity based loss function for joint regression and classification in AD diagnosis," NeuroImage, vol. 100, pp. 91–105, 2014. [26] F. Li, L. Tran, K.-H. Thung, S. Ji, D. Shen, and J. Li, "A robust deep model for improved classification of AD/MCI patients," IEEE J. Biomed. Health Inform., vol. 19, no. 5, pp. 1610–1616, 2015. [27] C. Zu, B. Jie, M. Liu, S. Chen, D. Shen, D. Zhang, A. D. N. Initiative et al., "Label-aligned multitask feature learning for multimodal classification of Alzheimers disease and mild cognitive impairment," Brain Imaging and Behavior, pp. 1–12, 2015. [28] A. Payan and G. Montana, "Predicting alzheimer's disease: a

International journal of imaging science and engineering
ISSN: 1934--9955 www.ijise.net
Vol-20 Issue-01 Mar 2025

neuroimaging study with 3D convolutional neural networks," arXiv:1502.02506 [cs.CV], 2015. [29] S. Liu, S. Liu, W. Cai, S. Pujol, R. Kikinis, and D. Feng, "Early diagnosis of Alzheimer's disease with deep learning," in Biomedical Imaging (ISBI), 2014 IEEE 11th International Symposium on. IEEE, 2014, pp. 1015–1018. [30] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," Proc. IEEE, vol. 86, no. 11, pp. 2278–2324, 1998. [31] E. Hosseini-Asl, J. Zurada, and O. Nasraoui, "Deep learning of part-based representation of data using sparse autoencoders with nonnegativity constraints," Neural Networks and Learning Systems, IEEE Transactions on, vol. PP, no. 99, pp. 1–13, 2015. [32] J. Masci, U. Meier, D. Cires¸an, and J. Schmidhuber, "Stacked convolutional auto-encoders for hierarchical feature extraction," in Artificial Neural Networks and Machine Learning–ICANN 2011. Springer, 2011, pp. 52–59. [33] A. Makhzani and B. Frey, "A winner-take-all method for training sparse convolutional autoencoders," arXiv:1409.2752 [cs.LG, cs.NE], 2014. [34] B. Leng, S. Guo, X. Zhang, and Z. Xiong, "3D object retrieval with stacked local convolutional autoencoder," Signal Processing, vol. 112, pp. 119 – 128, 2015. [35] X. Glorot and Y. Bengio, "Understanding the difficulty of training deep feedforward neural networks," in Proceedings of the International Conference on Artificial Intelligence and Statistics (AISTATS10). Society for Artificial Intelligence and Statistics, 2010, pp. 249–256. [36] S. Thrun, "Is learning the n-th thing any easier than learning the first?" Advances in Neural Information Processing Systems, pp. 640–646, 1996. [37] R. Caruana, "Multitask learning," Machine Learning, vol. 28, no. 1, pp. 41–75, 1997. [38] R. Raina, A. Y. Ng, and D. Koller, "Constructing informative priors using transfer learning," in Proceedings of the 23rd International Conference on Machine Learning. ACM, 2006, pp. 713–720. [39] J. Baxter, "A Bayesian/information theoretic model of learning to learn via multiple task sampling," Machine Learning, vol. 28, no. 1, pp. 7–39, 1997. [40] J. S. Bridle and S. Cox, "RecNorm: Simultaneous normalisation and classification applied to speech recognition." in Advances in Neural Information Processing Systems, 1990, pp. 234–240. [41] S. Ben-David, J. Blitzer, K. Crammer, A. Kulesza, F. Pereira, and J. W. Vaughan, "A theory of learning from different domains," Machine Learning, vol. 79, no. 1, pp. 151–175, 2009. [42] K. Crammer, M. Kearns, and J. Wortman, "Learning from multiple sources," The Journal of Machine Learning Research, vol. 9, pp. 1757–1774, 2008. [43] G. Mesnil, Y. Dauphin, X. Glorot, S. Rifai, Y. Bengio, I. J. Goodfellow, E. Lavoie, X. Muller, G. Desjardins, D. Warde-Farley et al., "Unsupervised and transfer learning challenge: a deep learning approach." ICML Unsupervised and Transfer Learning, vol. 27, pp. 97– 110, 2012. [44] X. Glorot, A. Bordes, and Y. Bengio, "Domain adaptation for large-scale sentiment classification: A deep learning approach," in Proceedings of the 28th International Conference on Machine Learning (ICML-11), 2011, pp. 513–520. [45] J. Weston, F. Ratle, H. Mobahi, and R. Collobert, "Deep learning via semi-supervised embedding," in Neural Networks: Tricks of the Trade. Springer, 2012, pp. 639–655. [46] M. D. Zeiler, "ADADELTA: An adaptive learning rate method," arXiv:1212.5701 [cs.LG], 2012. [47] F. Bastien, P. Lamblin, R. Pascanu, J. Bergstra, I. Goodfellow, A. Bergeron, N. Bouchard, D. Warde-Farley, and Y. Bengio, "Theano: new features and speed improvements," arXiv:1211.5590 [cs.SC], 2012. [48] L. Van der Maaten and G. Hinton, "Visualizing Data using t-SNE." Journal of Machine Learning Research, vol. 9, no. 11, 2008. [49] R. H. Fletcher, S. W. Fletcher, and G. S. Fletcher, Clinical epidemiology: the essentials. Lippincott Williams & Wilkins, 2012